Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

How to interpret the total number of SARS-CoV-2 infections

Published Online
April 8, 2022
https://doi.org/10.1016/S0140-6736(22)00629-8
See Online/Article
https://doi.org/10.1016/S0140-6736(22)00484-6

Counts of reported cases have been the key metric to monitor the COVID-19 pandemic. However, since the beginning, it has been clear that reported cases represent only a fraction of all SARS-CoV-2 infections.[1] In *The Lancet*, Ryan Barber and colleagues, writing on behalf of the Institute for Health Metrics and Evaluation, report a comprehensive set of global and location-specific estimates of daily and cumulative SARS-CoV-2 infections and the proportion of the population infected for 190 countries and territories up to Nov 14, 2021.[2] For this, the authors used a novel approach, combining data from reported cases and deaths, excess deaths attributable to COVID-19, hospitalisations, and seroprevalence surveys to produce more robust estimates in an attempt to minimise biases. According to Barber and colleagues' findings, a staggering number of people, 3·39 billion (95% uncertainty interval 3·08–3·63) or 43·9% (39·9–46·9) of the global population, are estimated to have been infected one or more times between March, 2020, and November, 2021. Remarkably, this was before the highly transmissible omicron (B.1.1.529) variant swept the globe. These estimates of total infections are wildly different from the number of reported cases, which stood at 254 million as of Nov 14, 2021.[3]

Barber and colleagues' study also highlights vast regional discrepancies, painting a very different picture from that provided by reported cases. From case reports, one would conclude that the highest cumulative incidence was observed in Europe and North America and the lowest in Africa. However, this study estimated that 70·5% (61·6–75·9) of the population in sub-Saharan Africa has been infected with SARS-CoV-2, compared with 30·9% (28·8–32·8) of the population in high-income North America. Underlying this apparent reversal of patterns are stark differences in case detection; fewer than 1% of infections were reported as cases in sub-Saharan Africa whereas nearly half were reported in high-income North America. It is crucial that this underreporting is considered when we compare the impact of the pandemic and the effectiveness of responses among nations.

It is also worth reflecting on the technical achievement in data integration that underpins these new estimates. Barber and colleagues were able to estimate cumulative infections at the national and subnational levels by integrating an array of data sources. Each individual dataset—cross-sectional serosurveys and time series of cases, hospitalisations, and deaths—has limited value and inherent bias on its own. Serosurveys are of highly variable quality, death reporting is incomplete,[4] and many outcomes are not reliably stratified by age or other key variables such as gender, race, and vaccination status. Despite the serious challenges in data integration on this scale and with this diversity of sources, it enables objective comparisons about the level of infection in a setting and can, for example, guide more optimal targeting of vaccines.

Although estimates of the proportion of the population ever infected provide insight into the cumulative impact and current phase of the epidemic in each location, we should be cautious not to conflate the proportion of the population ever infected with population-level immunity. The proportion ever infected, combined with vaccine coverage, has been proposed as a metric to evaluate whether we have reached sufficient population immunity to stop widespread community transmission. However, with new variants escaping immunity, immunity waning, and unequal distribution of vaccination, defining population-level immunity is not trivial.[5] Barber and colleagues' study estimated population immunity in the simplest way possible: by assuming that previously infected people were immune, vaccination was randomly distributed, and immunity did not wane. Tellingly, this metric did not inversely correlate with community transmission (ie, the time-varying reproductive number), showing that such a simple approach no longer provides an appropriate measure of population immunity. A more reliable measure would account for waning, boosting from multiple exposures, non-random vaccine uptake, different immune response across age groups, and cross-variant immunity.

As such, one could argue that the proportion of the population ever infected is no longer a meaningful metric of population immunity. However, the same data streams to infer cumulative incidence can be used to address more pressing epidemiological questions,

such as how severe are new variants? To what extent do the population's historical infections—in terms of timing and variants—protect against infection and severe disease of new variants? Relatedly, how do layers of vaccine-induced and virus-induced immunity combine to confer protection to the population? Perhaps most importantly at this moment in the pandemic, we need to identify the sub-populations that remain susceptible to severe disease and death. Serosurveys combined with morbidity and mortality surveillance and detailed monitoring of vaccine coverage are essential to identify the groups lacking immunity from vaccination or previous infection.[5–7] Integrating data enables the kinds of insights offered by Barber and colleagues to inform the next phase of the pandemic response, and we should sustain this effort.

BL reports personal consulting fees from Epidemiologic Research & Methods, outside the submitted work. KS declares no competing interests.

*Kayoko Shioda, *Ben Lopman*
blopman@emory.edu

Department of Epidemiology and Gangarosa Department of Environmental Health, Rollins School of Public Health, Emory University, Atlanta, GA 30322, USA

1 Russell TW, Golding N, Hellewell J, et al. Reconstructing the early global dynamics of under-ascertained COVID-19 cases and infections. *BMC Med* 2020; **18:** 332.
2 COVID-19 Cumulative Infection Collaborators. Estimating global, regional, and national daily and cumulative infections with SARS-CoV-2 through Nov 14, 2021: a statistical analysis. *Lancet* 2022; published online April 8. https://doi.org/10.1016/S0140-6736(22)00484-6.
3 Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. *Lancet Infect Dis* 2020; **20:** 533–34.
4 COVID-19 Excess Mortality Collaborators. Estimating excess mortality due to the COVID-19 pandemic: a systematic analysis of COVID-19-related mortality, 2020–21. *Lancet* 2022; published online March 10. https://doi.org/10.1016/S0140-6736(21)02796-3.
5 Lopman BA, Shioda K, Nguyen Q, et al. A framework for monitoring population immunity to SARS-CoV-2. *Ann Epidemiol* 2021; **63:** 75–78.
6 Mina MJ, Metcalf CJE, McDermott AB, Douek DC, Farrar J, Grenfell BT. A global Immunological observatory to meet a time of pandemics. *eLife* 2020; **9:** e58989.
7 Elliott P, Bodinier B, Eales O, et al. Rapid increase in omicron infections in England during December 2021: REACT-1 study. *Science* 2022; **375:** 1406–11.